UNC
ESHELMAN
SCHOOL OF PHARMACY

MML
UNC.EDU

Computer-Aided Design of Nanomaterials with the Desired Bioactivity and Safety Profiles

Denis Fourches and Alexander Tropsha
University of North Carolina at Chapel Hill, USA

in collaboration with ***Bing Yan***
St Jude Children Research Hospital, USA

Outline: From data to models to computer-aided design

MML
UNC.EDU

1. Data Accumulation: Need to establish an integrated database of nanomaterials' structure and properties

2. Data modeling and nano-informatics: Need to develop new descriptors for nanomaterials

3. Application of the Quantitative Structure Activity Relationships (QSAR) approach to modeling of nanomaterials

4. Proof-of-concept: Computer aided design of carbon nanotubes with the desired bioactivity and safety profiles

- **~1000** manufacturer-identified nanotechnology-based consumer products currently on the market
  (Woodrow Wilson International Center for Scholars, 2008).
  Recent boost due in part to the application of combinatorial chemistry and high throughput screening to design novel Manufactured NanoParticles (**MNPs**)

- Growing public concern about the safety of MNPs since it has been demonstrated that MNPs intended for industrial applications could cause toxic effects in humans.
  Myllynen, P. Nat. Nanotechnol., 2009, 4, 795-796.
  Kipen et al. Am J Physiol Lung Cell Mol Physiol, 2005, 289, 696-697.

Experimental, toxicological testing of MNPs is costly and time-consuming and could thus restrict the development of newly designed particles.

Development of predictive « *in silico* » approaches

Main questions yet to be solved

Which MNPs are the most toxic?

What are the three most meaningful assays to determine MNP-induced toxicological effects?

What effects are already known for this type of particles?

Compared to this given organic [illegible] is a particular MNP more or less toxic at the same [illegible]

Which are the nanomaterials to be tested first?

Are there correlations between the properties of MNPs and their biological characteristics?

BUT WHERE ARE THE DATA?

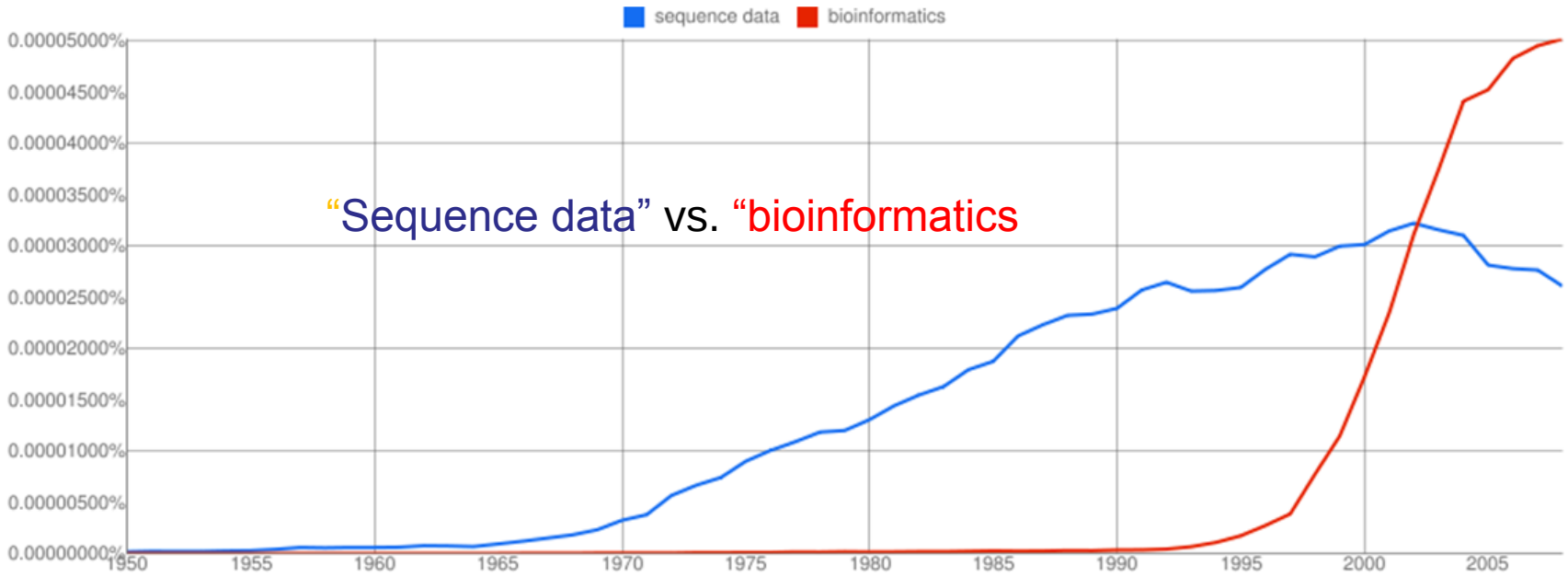
At what dose do most MNPs start to be inducing toxic effects?

The importance of the DATA to enable any informatics-dependent discipline: *bioinformatics example*

MML UNC.EDU

Google labs Books Ngram Viewer

Graph these **case-sensitive** comma-separated phrases: sequence data,bioinformatics
between 1950 and 2008 from the corpus English with smoothing of 3 .
Search lots of books



"Sequence data" vs. "bioinformatics

Search in Google Books:

The importance of the DATA to enable any informatics-dependent discipline: *cheminformatics example*

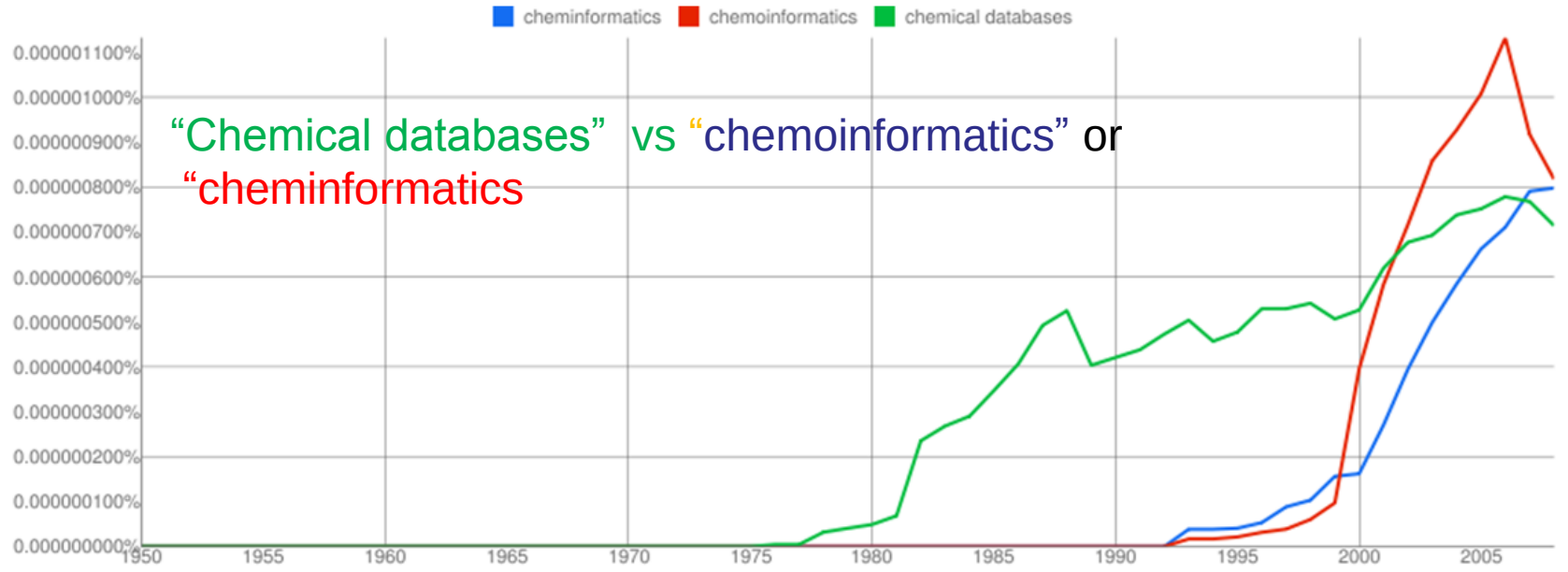
MML UNC.EDU

Google labs Books Ngram Viewer

Graph these **case-sensitive** comma-separated phrases: cheminformatics,chemoinformatics,chemical databases

between 1950 and 2008 from the corpus English with smoothing of 3 .

Search lots of books



Search in Google Books:

Done

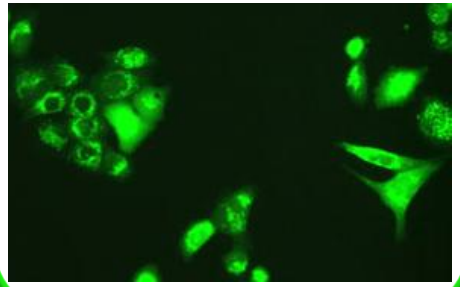
Internet | Protected Mode: C

The importance of the DATA to enable any informatics-dependent discipline: *nanotoxicoinformatics example*

- So far, three-four publications with enough data to build models
- Three-four computational modeling papers
- It is hard to collect and compile the data

Challenges of data integration: an example of Lung A549 cells.

Lung adenocarcinoma A549 cells



Pulskamp et al., Toxicol. Lett., 2007, 168, 58-74.

Several carbon MNPs (multi-walled, single-walled, carbon black, quartz) increased Reactive Oxygen Species (ROS) and decreased mitochondrial membrane potential in a dose- and time-dependent manner in rat macrophages and human **A549** lung cells.

Liu et al., Nanotechnology, 2010, 21, 315106.

The authors demonstrated the efficiency for lung cancer treatment of nanodiamond NPs carrying paclitaxel on their surface: these NPs were found *(i)* to reduce the **A549** cell viability *in vitro* by inducing both mitotic arrest and apoptosis, and *(ii)* blocked the tumor growth in mice.

Tahara et al., Int. J. Pharm., 2009, 382, 198-204.

The A549 cell uptake of chitosan-modified PLGA nanospheres is time-, temperature-, and concentration-dependent, regulated by clathrin-mediated endocytosis. Low cytotoxicity was reported for these modified, surface decorated nanospheres, suggesting them as preferable drug carriers for **A549** cells.

Deng et al., Nanotoxicology., 2010, 4, 186-195.

Foldbjerg et al., Arch. Toxicol., 2010, In Press.

PVP coated silver nanoparticles were reported to induce ROS and damage DNA in **A549** cells depending on their doses, as well as increase gap junctional intercellular communication.

Data depositories

Studies on MNP of different core structure, size, shape, and with various surface modifications have been reported but all published data are diverse, non-searchable, and spread among numerous sources of information.

- Lack of centralized data repository
- Limits our capability to develop predictive tools to assess nanotoxicity in advance of manufacturing
- Severely limits the design of novel nanomaterials that are environmentally benign and safe for human exposure

Specific Aim 1 *SRC White Paper 2011 / Tropsha group*

To compile, curate and organize a specialized database incorporating all existing information on MNP including their physical/chemical properties and associated biological data emerging both from SRC research teams and scientific literature.

- Will facilitate research collaboration and data sharing between research teams;
- Will enable computational modeling by providing larger sets of integrated and curated data;
- Will highly benefit both experimentalists and modelers by enabling easily accessible, efficient data storage and in-depth analysis/modeling of all reported experiments.

Data sharing/storing format for nanomaterials

Nanotechnology data sharing and standards >

nano-TAB

nano-TAB is a tab-delimited spreadsheet type of format facilitating the submission and exchange of data pertaining to nanomaterials and their characterizations (physico-chemical, *in vitro*, and *in vivo*). nano-TAB is based on existing standards developed by the European Bioinformatics Institute (EBI) and the Investigation/Study/Assay (ISA-TAB) file format, which represents a variety of assays and technology types. The nano-TAB specification leverages ISA-TAB files describing investigations, studies, and assays and provides extensions to support nanomaterial structural information and concepts on nanotechnology assay measurements defined in the NanoParticle Ontology (NPO).

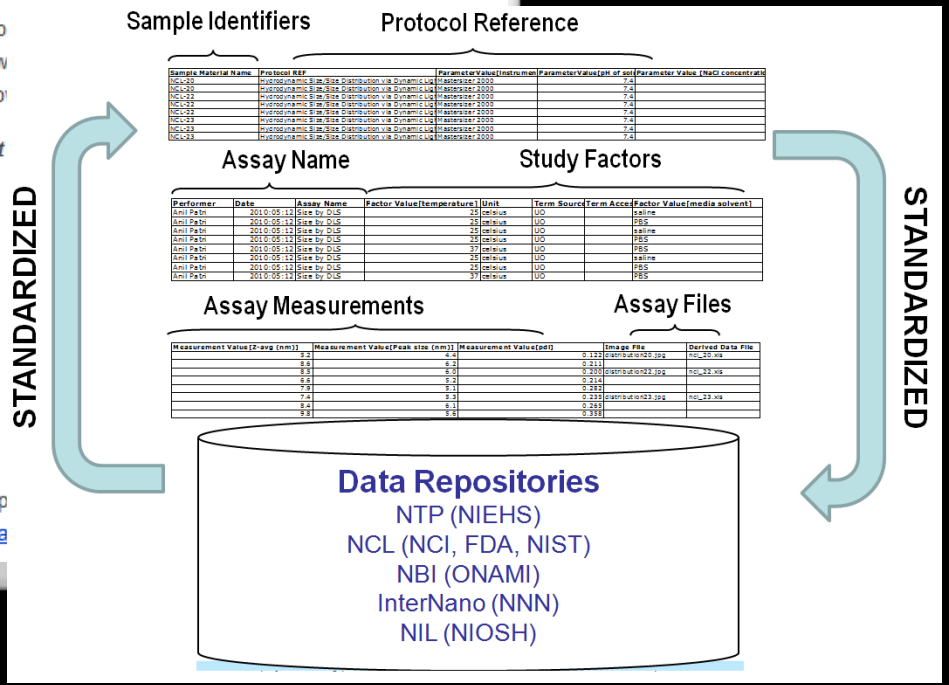
The goals of nano-TAB are to:

- Enable the submission and exchange of nanomaterials to/from nanotechnology resources
- Empower organizations to adopt standards for representing data in nanotechnology publications, and
- Provide researchers with guidelines for representing nanomaterials and characterizations to achieve cross-material comparison

The nano-TAB project is a sub-group of the caBIG Nano WG and has been assisted in clo
Representation (IR) Working Group (WG). The nano-TAB project is a collaborative effort w
registered ASTM Work Item WK28974. Links to key nano-TAB artifacts are provided belo

- ***nano-TAB Overview Presentation*** ***(this is the best starting place to learn about***
- nano-TAB Specification [DRAFT]
- nano-TAB Template
- nano-TAB Template Glossary
- nano-TAB Example Files:
  - NCL Dendrimer MRI Constrast Agent
  - NBI AU Nanoparticle
- nano-ML (nano-TAB XML Representation)
  - nano-ML XML Schema
  - nano-ML Example File

nano-TAB project meeting presentations, minutes and other files on the nano-TAB develop
http://gforge.nci.nih.gov/docman/index.php?group_id=69&selected_doc_group_id=5653&la



https://sites.google.com/site/cabignanowg/data-sharing-and-nanotechnology-standards/nanotab

Data sharing/storing format for nanomaterials

MML UNC.EDU

Study File

STUDY	
Study Identifier	NCL20061224-5178
Study Title	Size by DLS
Study Submission Date	Dec-06
Study Public Release Date	Dec-06
Study Description	Dynamic light scattering (DLS) technique was used to measure the hydrodynamic size of this dendritic nanomaterial. The effects of sample concentration, buffer, and temperature on the hydrodynamic size (stability) also were measured. Purity was analyzed by HPLC and Capillary Electrophoresis. MALDI-TOF Mass spectrometry was used to obtain the molecular weight information and to determine the purity, existence of dimers, trimers, and trailing generations in the sample. Gradeless quantitation, which is important to determine the reactivity as a MRI contrast agent, was carried out by Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES). Fractionation methods such as Size Exclusion Chromatography (SEC) and Asymmetric-flow Field Flow Fractionation (AFFF) were used to determine the molecular weight information as well as purity. Finally, a 3T clinical MRI machine was used to obtain relaxivity measurements on this sample to compare with free Magnevist®. No significant relaxivity change was observed upon association of Magnevist® with the dendrimer
Study Disease	
Study Disease Term Accession Number	
Study Disease Term Source REF	
Study Outcome	For NCL22, the size is slightly larger when dispersed in saline compared to PBS. In PBS, the size is independent of temperature. This is in contrast to NCL23, which is larger in PBS than in saline. NCL23 also shows temperature dependence, as its size decreases slightly with increased temperature in PBS. Finally, NCL20 is larger when dispersed in PBS compared to saline.
STUDY SAMPLES	
Study Sample File Name	m_NCL-21.txt; m_NCL-22.txt; m_NCL-23.txt
Study Sample File URI	
Study Sample File Type	material sample; material sample; material sample
Study Sample File Type Term Accession Number	
Study Sample File Type Term Source REF	
Study Sample File Version	
Study Sample File Description	
STUDY DESIGN DESCRIPTORS	
Study Design Type	comparison
Study Design Type Term Accession Number	
Study Design Type Term Source REF	
STUDY PUBLICATIONS	
Study PubMed ID	
Study Publication DOI	
Study Publication Author list	
Study Publication Title	
Study Publication Status	
Study Publication Status Term Accession Number	
Study Publication Status Term Source REF	

Study Description

Sample Files

Design

Publication

Investigation File

ONTOLOGY SOURCE REFERENCE		
Term Source Name	MO	MO
Term Source File		
Term Source Version	v1.0	v1.0
Term Source Description	MGED Ontology	Nanoparticle Ontology
INVESTIGATION		
Investigation Identifier	NCL20061224	
Investigation Title	Dendrimer-based MRI Contrast Agents	
Investigation Description	The objective of the Dendritic Nanotechnologies, Inc. – NCL collaboration is to characterize a PAMAM dendrimer with an associated gadolinium chelate MRI contrast agent. The nanomaterials submitted for testing at the NCL were (NCL20) G4 bis (hydroxyl) terminated PAMAM dendrimer, (NCL21) G4 pyrrolidinone terminated PAMAM dendrimer, (NCL22) G4.5 COOH terminated PAMAM dendrimer, (NCL23) G4.5 COONa terminated PAMAM dendrimer-Hogen/a6® complex, and (NCL24) G4 pyrrolidinone terminated PAMAM dendrimer-Hogen/a6® complex. Commercially available Hogen/a6® (NCL24) was used as a control. NCL studies addressed in this registry can be divided into three main categories: physicochemical characterization; immunotoxicology; in vitro toxicology	
Investigation Disease		
Investigation Disease Term Accession Number		
Investigation Disease Term Source REF		
Investigation Outcome		
Investigation Submission Date	12/1/2006	
Investigation Public Release Date	12/1/2006	
INVESTIGATION PUBLICATIONS		
Investigation PubMed ID	12093846	
Investigation Publication DOI	10.2217/17435889.2.6.780	
Investigation Publication Author list	Heidi JB; Dobrovolskaia MA; Patri AK; McNeil SE	
Investigation Publication Title	Characterization of nanoparticles for therapeutics	
Investigation Publication Status	peer reviewed	
Investigation Publication Status Term Accession Number		
Investigation Publication Status Term Source REF	NPO	
INVESTIGATION CONTACTS		
Investigation Person Last Name	McNeil	Patri
Investigation Person First Name	Scott	Anil
Investigation Person Mid Initials		
Investigation Person Email	mcneils@mail.nih.gov	patria@mail.nih.gov
Investigation Person Phone	3018469939	3018465237
Investigation Person Fax		
Investigation Person Address	NCI-FCRDC 1050 Boyles Street, Frederick, MD 21702	NCI-FCRDC 1050 Boyles Street, F
Investigation Person Affiliation	Nanotechnology Characterization Laboratory	Nanotechnology Characte
Investigation Person Roles	Investigator	Investigator
Investigation Person Roles Term Accession Number		
Investigation Person Roles Term Source REF	MO	MO

Ontology References

Investigation Description

Publications

Contacts

Material File

Material Component Identifier

Material Component Characteristics

Material Component Files

Nano-TAB

Assay File

Sample Identifiers

Protocol Reference

Material Sample Name	Protocol REF	ParameterValue[pH of solutio	Parameter Value [NaCl concentration	Performer	Date
NCL-20-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-20-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-22-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-22-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-22-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-23-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-23-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12
NCL-23-F1	Hydrodynamic Size/Size Distribution via Dynamic Ligl	7.4		Anil Patri	2010:05:12

Assay Name

Study Factors

Assay Name	Factor Value[temperature]	Unit	Term Source	Term Acces	Factor Value[media solvent]	Measurement Value[size]	Statistic	Unit	Term Source
Size by DLS	25	celsius	UO		saline	5.2	z-average	nm	UO
Size by DLS	25	celsius	UO		PBS	8.6	z-average	nm	UO
Size by DLS	25	celsius	UO		saline	8.5	z-average	nm	UO
Size by DLS	25	celsius	UO		PBS	6.6	z-average	nm	UO
Size by DLS	37	celsius	UO		PBS	7.9	z-average	nm	UO
Size by DLS	25	celsius	UO		saline	7.4	z-average	nm	UO
Size by DLS	25	celsius	UO		PBS	8.4	z-average	nm	UO
Size by DLS	37	celsius	UO		PBS	9.8	z-average	nm	UO

Assay Measurements

Assay Files

Term Acce	Measurement Value[Peak size]	Unit	Term Source	Term Acce	Measurement Value[pdI]	Image File	Derived Data File
	4.4	nm	UO		0.122	distribution20.jpg	ncl_20.xls
	6.2	nm	UO		0.211		
	6.0	nm	UO		0.200	distribution22.jpg	ncl_22.xls
	5.2	nm	UO		0.214		
	5.1	nm	UO		0.282		
	5.3	nm	UO		0.235	distribution23.jpg	ncl_23.xls
	6.1	nm	UO		0.265		
	5.6	nm	UO		0.358		

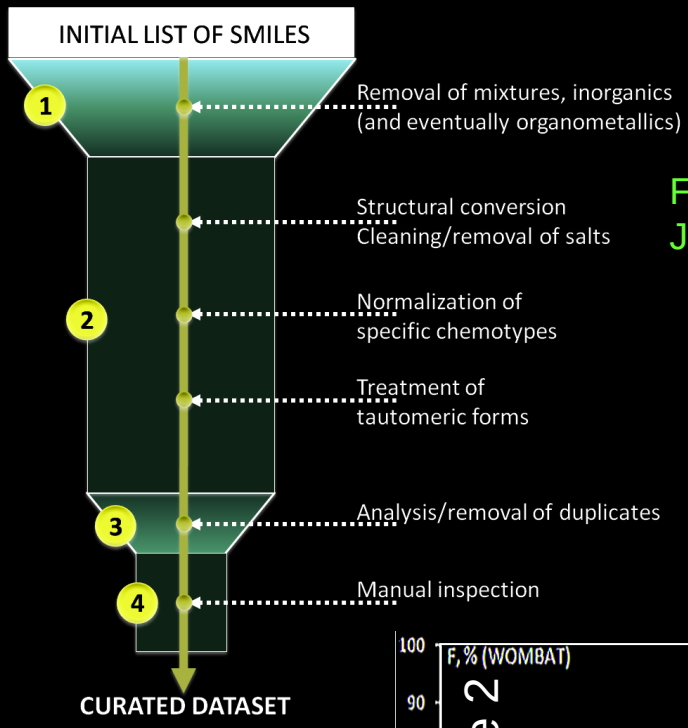
caBIG

Structure File (Optional)

Atom

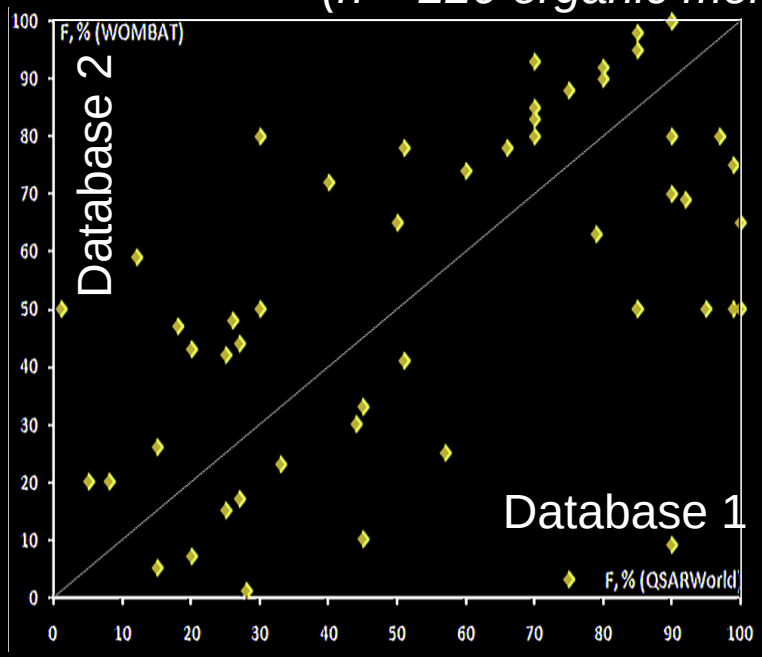
Header

Solving chemical and biological data curation issues



Fourches, D.; Muratov, E.; Tropsha, A.
J. Chem. Inf. Model. 2010, 50, 1189-1204.

Bioavailability data
(n = 220 organic molecules)



Similar Issues With Nanoparticles

Outline: From data to models to computer-aided design

MML
UNC.EDU

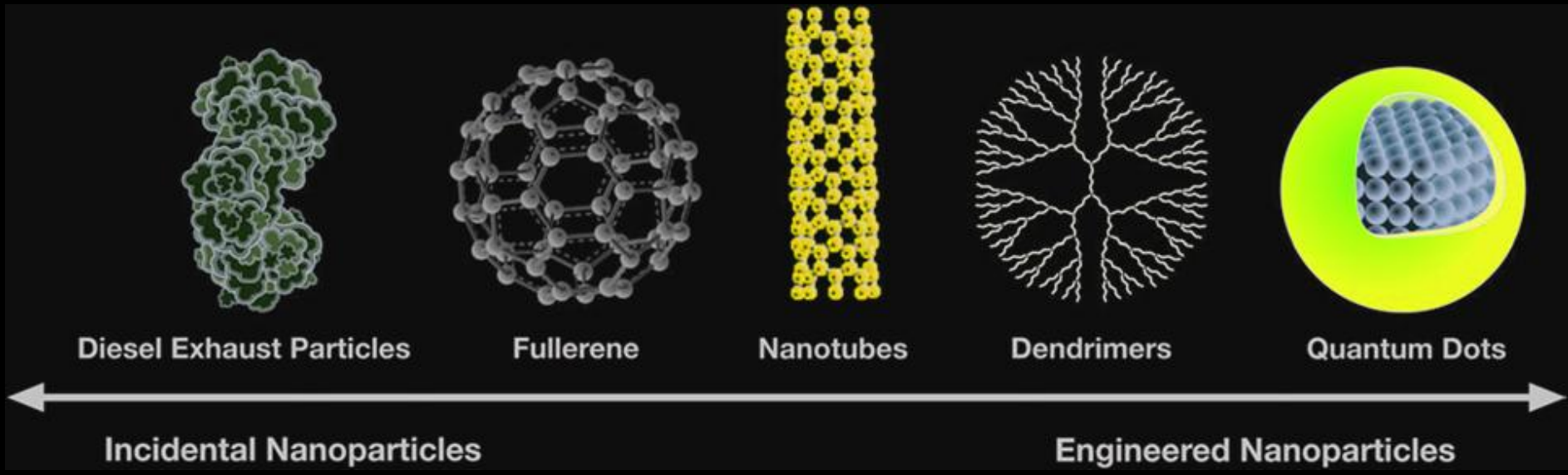
1. Data Accumulation: Need to establish an integrated database of nanomaterials' structure and properties

2. Data modeling and nano-informatics: Need to develop new descriptors for nanomaterials

3. Application of the Quantitative Structure Activity Relationships (QSAR) approach to modeling of nanomaterials

4. Proof-of-concept: Computer aided design of carbon nanotubes with the desired bioactivity and safety profiles

Difficulties in Modeling of Nanomaterials



S. Stern and S. McNeil, Toxicological Sciences, 101(1), 4-21, 2008.

- NP structures are very diverse → a real challenge to develop quantitative parameters (***descriptors***) of MNPs.
- Systematic physico-chemical, geometrical, structural and biological studies of NPs are nearly absent.
- Computational modeling of nanoparticles is only beginning to emerge; best if done in collaboration with experimental scientists.

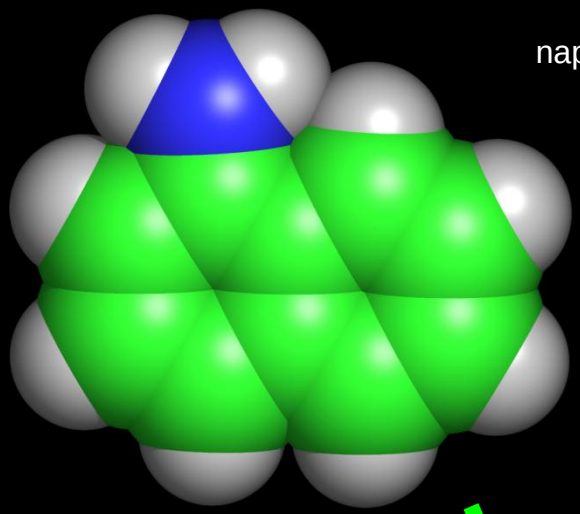
Simplest view of QNAR progression

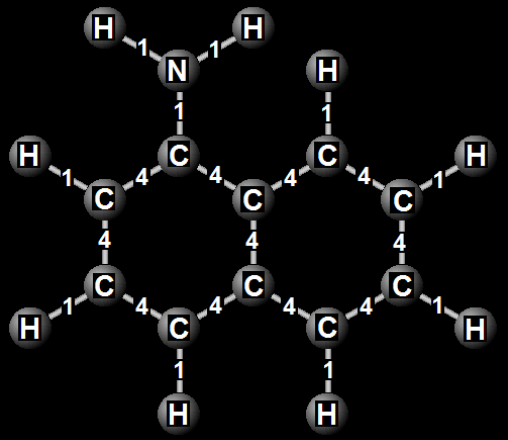
- Experimental Data *= pain*
  - Structure
  - Activity
- Validated models of data
  - Descriptors
  - Statistical/machine learning techniques
- Imputed data
- Experimentally confirmed predictions
- Reliable models to enable decision support (both in research and regulations) *= gain*

Structure representation in cheminformatics

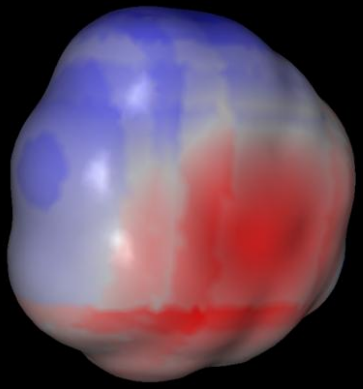
MML
UNC.EDU

naphtalen-1-amine

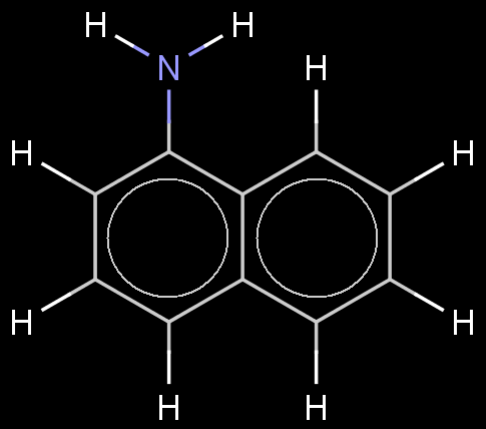




Viewed by computers



Viewed by another molecule

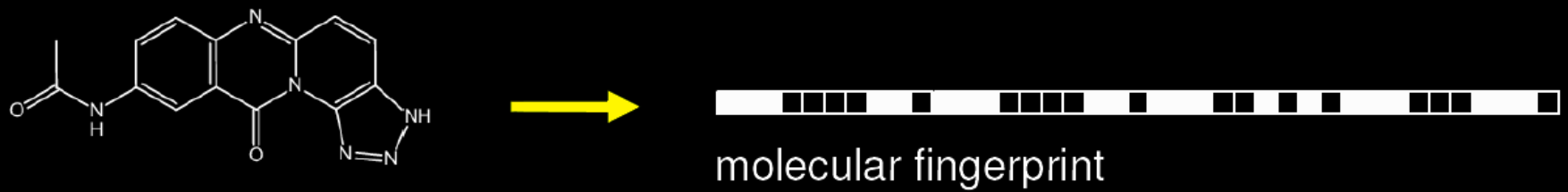


Viewed by chemists

Compounds are represented by a matrix of molecular descriptors

Samples (Compounds)	Variables (descriptors)			
	X_1	X_2	...	X_m
1	X_{11}	X_{12}	...	X_{1m}
2	X_{21}	X_{22}	...	X_{2m}
...	...	...	...	...
n	X_{n1}	X_{n2}	...	X_{nm}

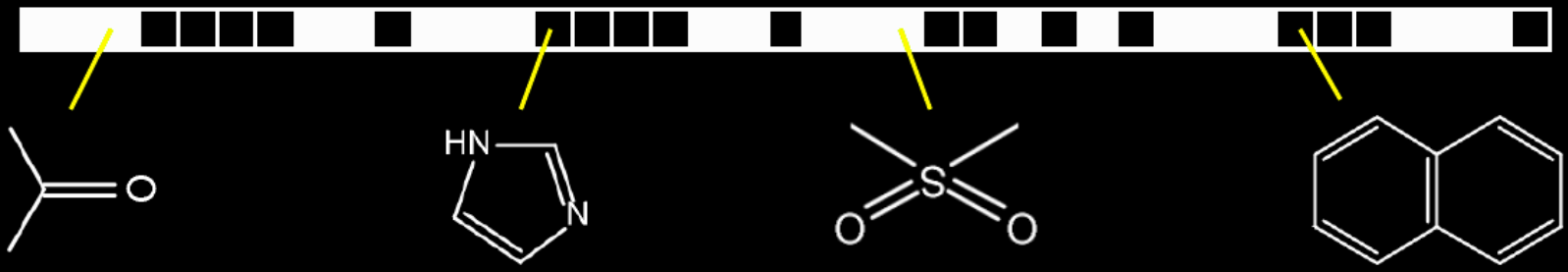
Molecular fingerprints - bit string encodings of structural features and/or calculated molecular properties.



Molecular Fingerprints

- 2D Fragment-based, keyed fingerprints: each bit position monitors the presence or absence of structural fragments

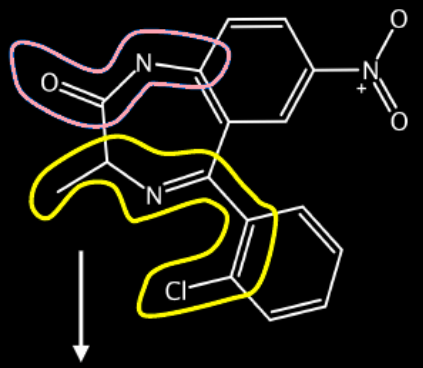
  MACCS (166 bits), BCI (e.g. 1,052 bits)



- 2D Hashed designs:

  Map different features (e.g. connectivity pathways) to overlapping bit segments

  Daylight (usually 2,048 bits)

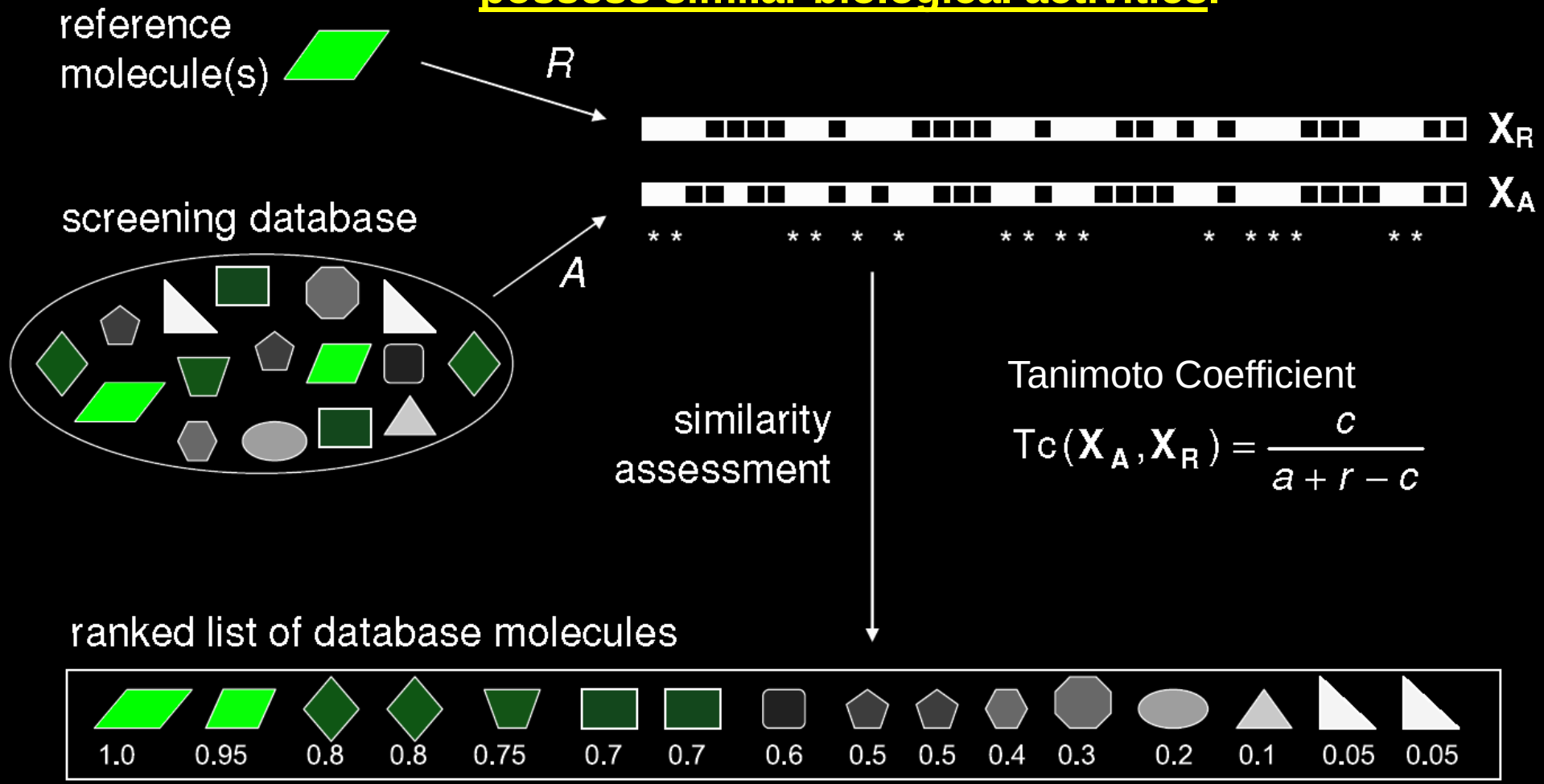


Pathways: O=CNC, ..., CCN=CC=CCl

From J. Bajorath, SSS Cheminformatics, Obernai 2008

Similarity Search

Similarity searching using fingerprint representations of molecules is one of the most widely used approaches for chemical database mining: it assumes that **similar compounds possess similar biological activities**.



From J. Bajorath, SSS Cheminformatics, Obernai 2008

MML
UNC.EDU

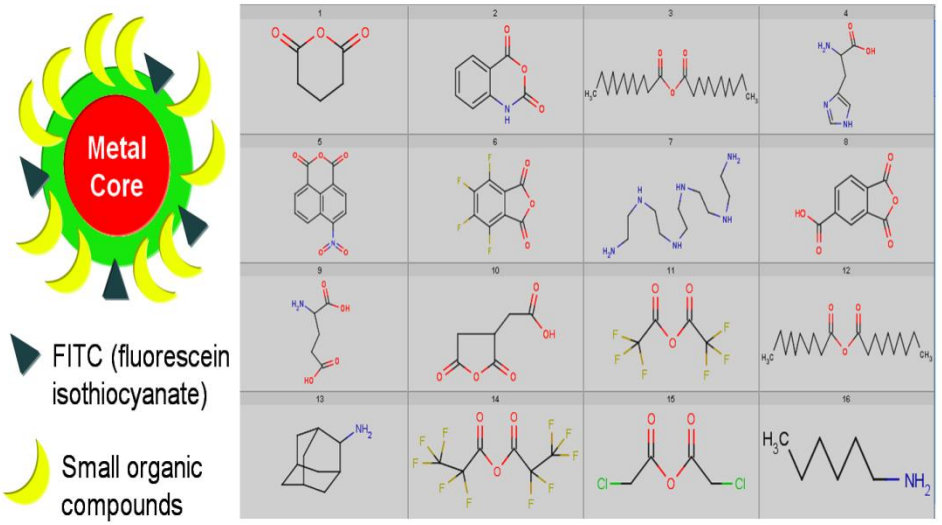
NANOPARTICLES

DESCRIPTORS

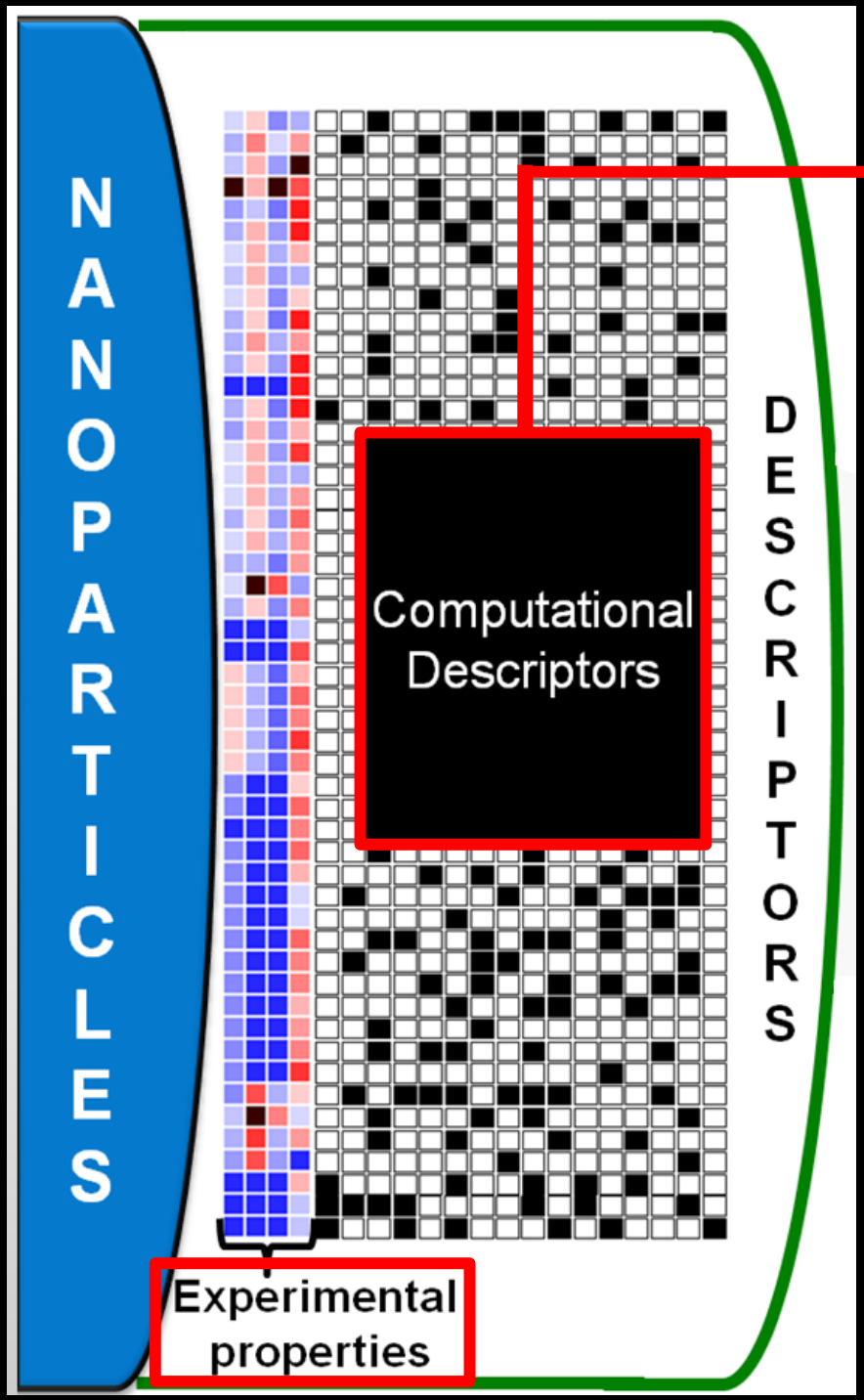
Computational Descriptors

Experimental properties

ALL PARTICLES HAVE THE SAME CORE BUT DIFFERENT SURFACE MODIFIERS



Classical molecular descriptors (e.g., Dragon, MOE, SiRMS) can be computed for a single molecule that represents the surface of a particular nanoparticle.



ALL PARTICLES HAVE DIFFERENT CORES/ARCHITECTURES

MNP	CLIO	PNP	MION	QD	Feridex IV	Ferrum Hausmann
#. particle	23	19	4	3	1	1

Shaw et al., PNAS, 2008, 105, 7387-7392

If no available three-dimensional structures, only constitutional descriptors (e.g., number of metal atoms, presence/ absence of coated dextran) are computationally accessible.

Need of developing new descriptors

Development of Quantum-Mechanics based fingerprints for thousands of nanomaterials in collaboration with Dr. Stefano Curtarolo *(Duke University)*

Outline: From data to models to computer-aided design

MML
UNC.EDU

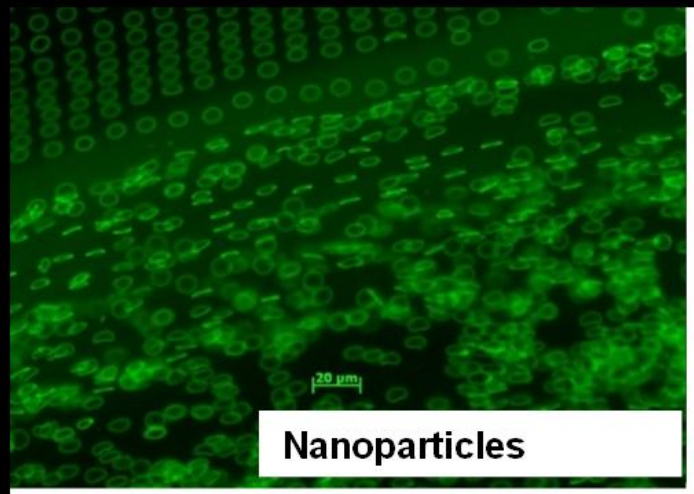
1.Data Accumulation: Need to establish an integrated database of nanomaterials' structure and properties

2. Data modeling and nano-informatics: Need to develop new descriptors for nanomaterials

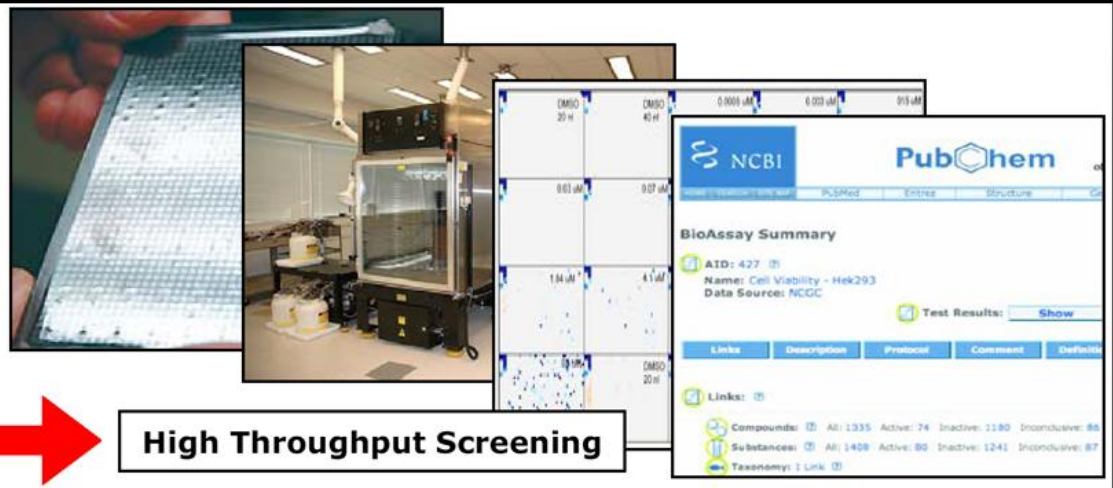
3. Application of the Quantitative Structure Activity Relationships (QSAR) approach to modeling of nanomaterials

4. Proof-of-concept: Computer aided design of carbon nanotubes with the desired bioactivity and safety profiles

Workflow for MNP risk assessment



Nanoparticles

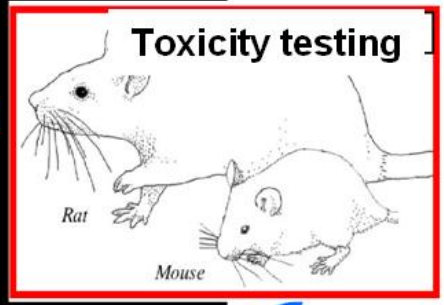


High Throughput Screening

Molecular weight, compositions and geometrical parameters, physico-chemical properties (acidic, basic, neutral, amphi- or lipophilic etc.)

Molecular properties

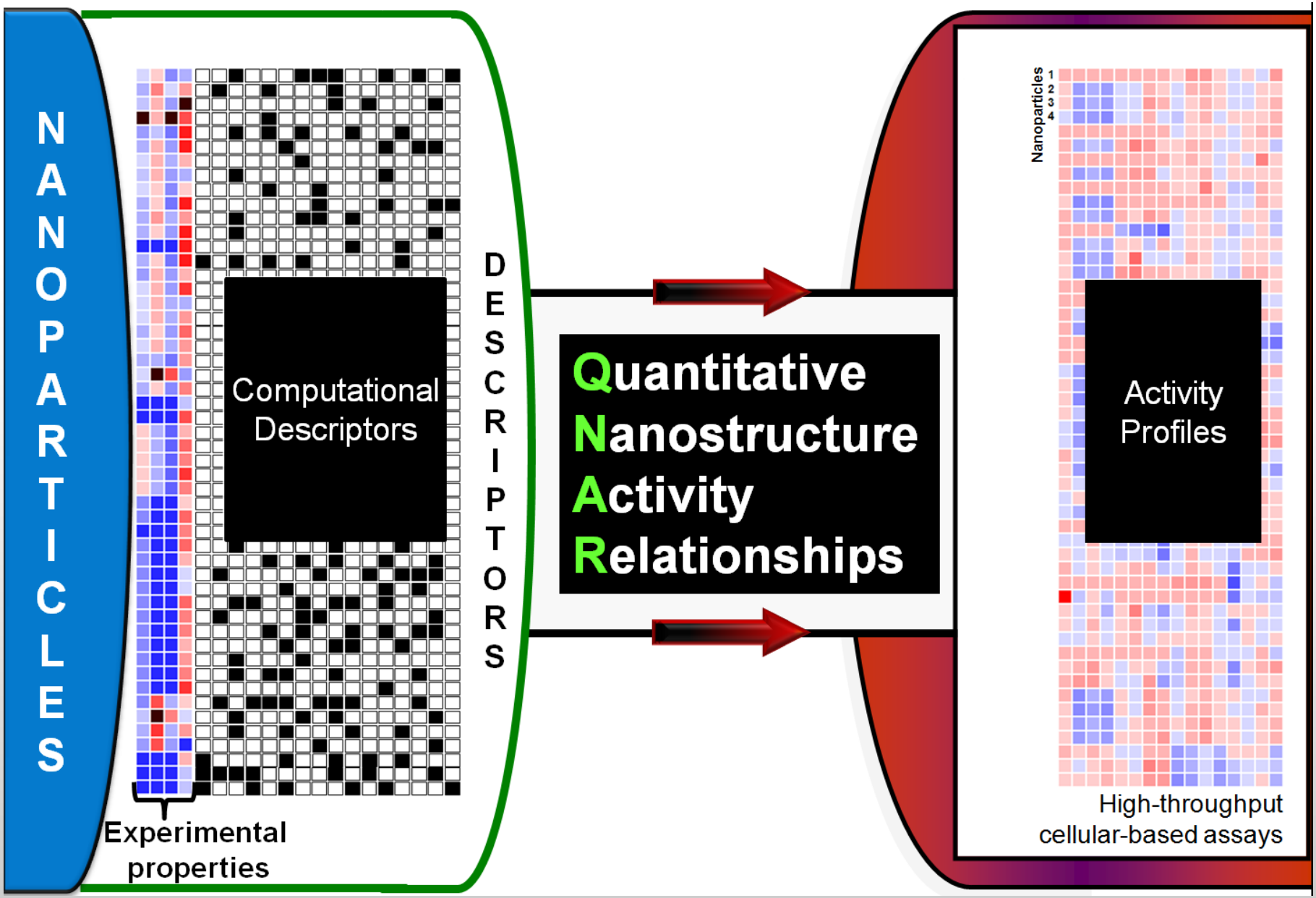
Human health risk



Biological Descriptors

Chemical Descriptors

"in silico" modeling



Fourches D, Pu D, Tropsha A. *Comb Chem High Throughput Screen*. 2011 Jan 26. [Epub ahead of print]

ACS Publications
MOST TRUSTED. MOST CITED. MOST READ.
ACS NANO
2010 Oct 26;4(10): 5703-12.

MML UNC.EDU

ARTICLE

Quantitative Nanostructure—Activity Relationship Modeling

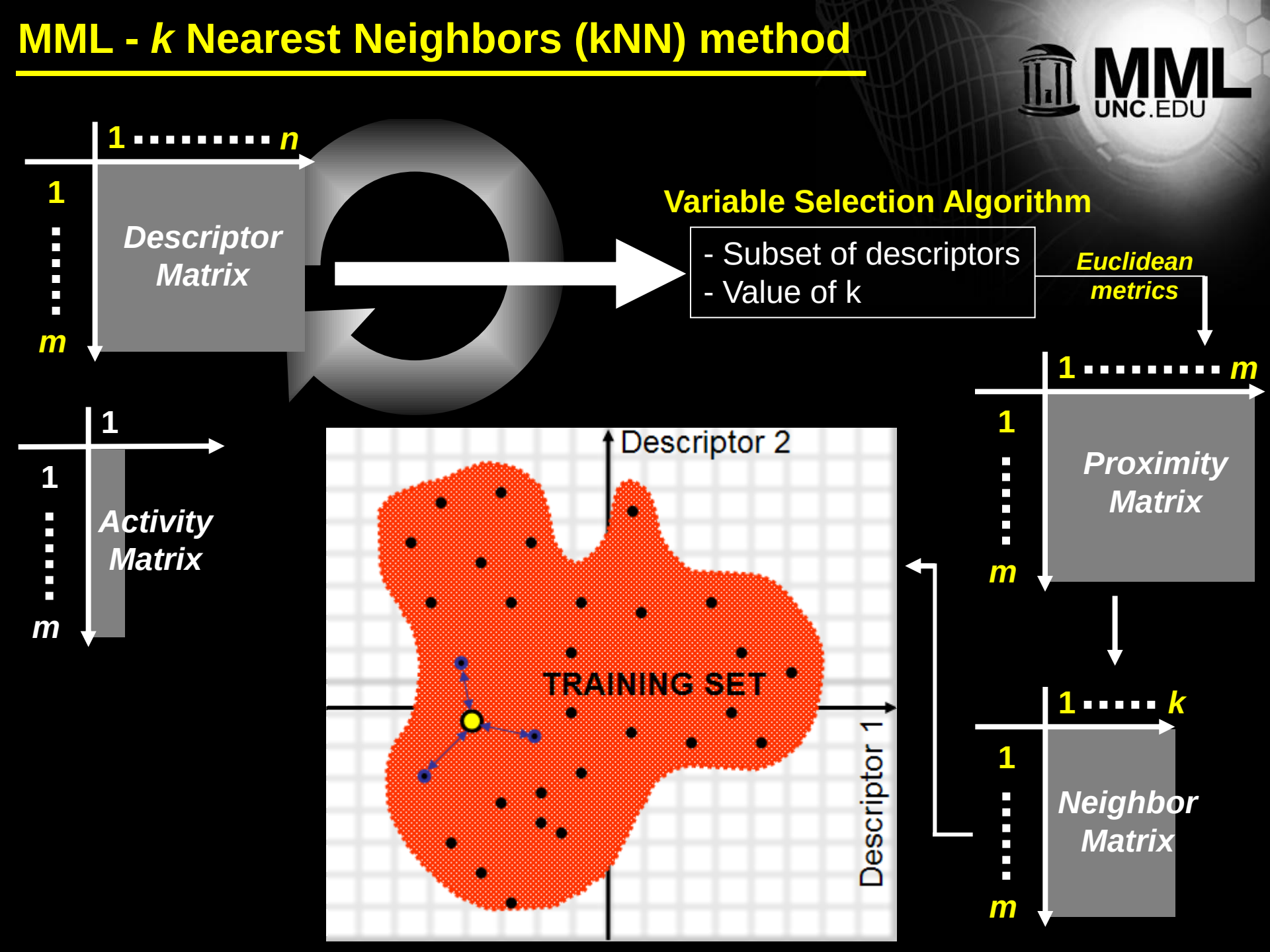
Denis Fourches,[†] Dongqiuye Pu,[†] Carlos Tassa,[‡] Ralph Weissleder,[‡] Stanley Y. Shaw,[‡] Russell J. Mumper,[§] and Alexander Tropsha[†,*]

[†]Laboratory of Molecular Modeling, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599, [‡]Center for Systems Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, and [§]Center for Nanotechnology in Drug Delivery, UNC Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599

ABSTRACT Evaluation of biological effects, both desired and undesired, caused by manufactured nanoparticles (MNPs) is of critical importance for nanotechnology. Experimental studies, especially toxicological, are time-consuming, costly, and often impractical, calling for the development of efficient computational approaches capable of predicting biological effects of MNPs. To this end, we have investigated the potential of cheminformatics methods such as quantitative structure—activity relationship (QSAR) modeling to establish statistically significant relationships between measured biological activity profiles of MNPs and their physical, chemical, and geometrical properties, either measured experimentally or computed from the structure of MNPs. To reflect the context of the study, we termed our approach quantitative nanostructure—activity relationship (QNAR) modeling. We have employed two representative sets of MNPs studied recently using *in vitro* cell-based assays: (i) 51 various MNPs with diverse metal cores (*Proc. Natl. Acad. Sci.* 2008, *105*, 7387—7392) and (ii) 109 MNPs

More than 1000 manufacturer-identified, nanotechnology-based consumer products are currently available on the market (The Woodrow Wilson International Center for Scholars, 2010). A growing fraction of them represent green products intended to achieve efficient and less polluting energy sources.[1] However, some manufactured nanoparticles (MNPs) intended for industrial applications may cause toxic effects in humans,[2−4] and public concern about the safety of

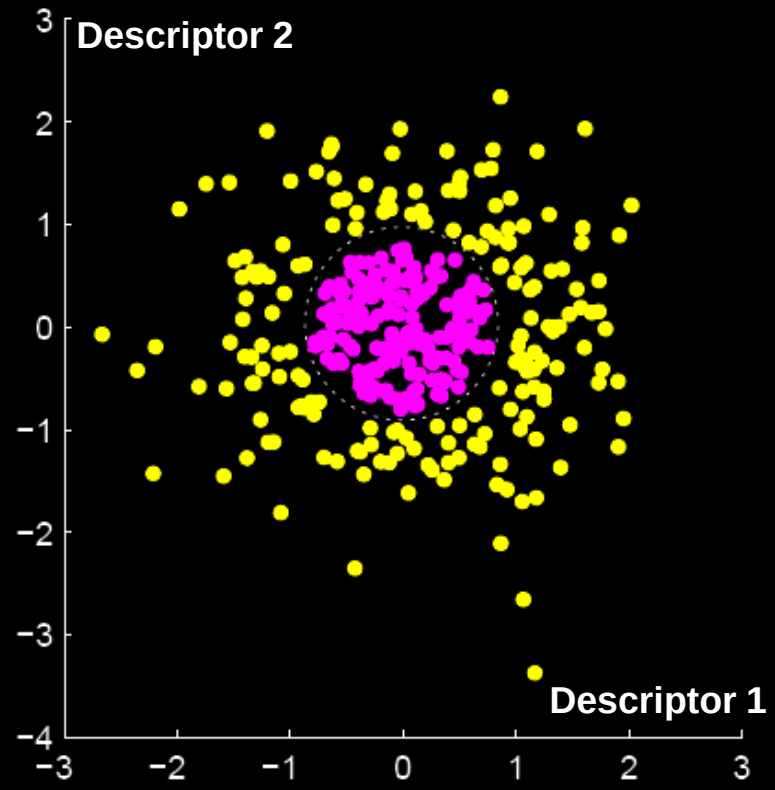
MML - *k* Nearest Neighbors (kNN) method



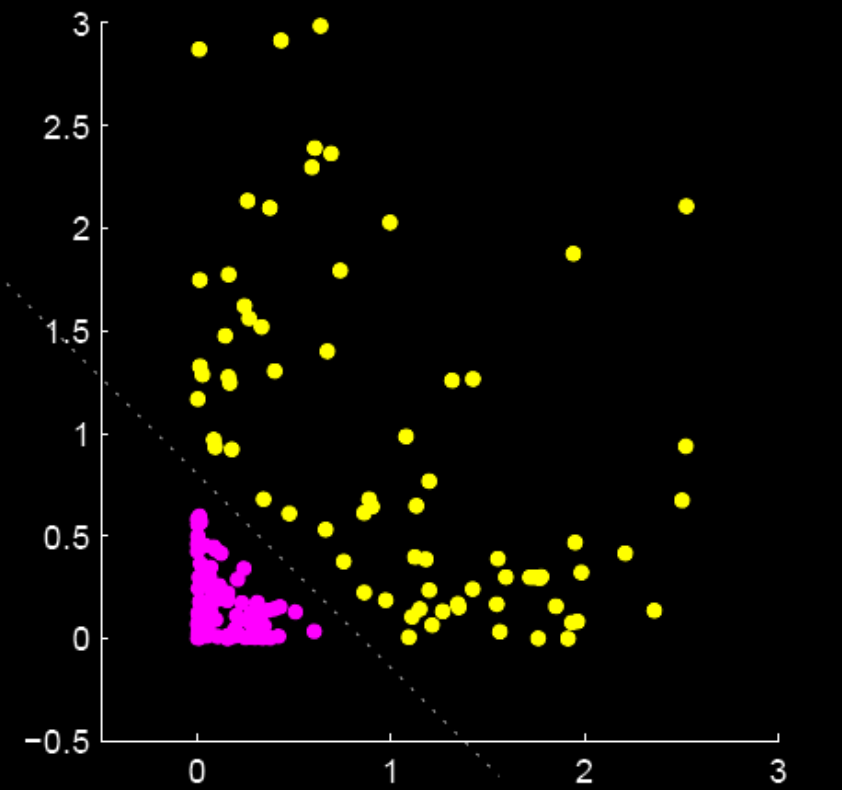
Support Vector Machine (SVM)

Introduced by Vapnik (1995), the SVM approach identifies the best linear separation between two classes of data. In a multidimensional descriptor space, such separation is realized by a hyperplane leading to the best linear segregation between data in the feature space.

What is the feature space ?



Kernel Function (Ex: x^2, y^2)



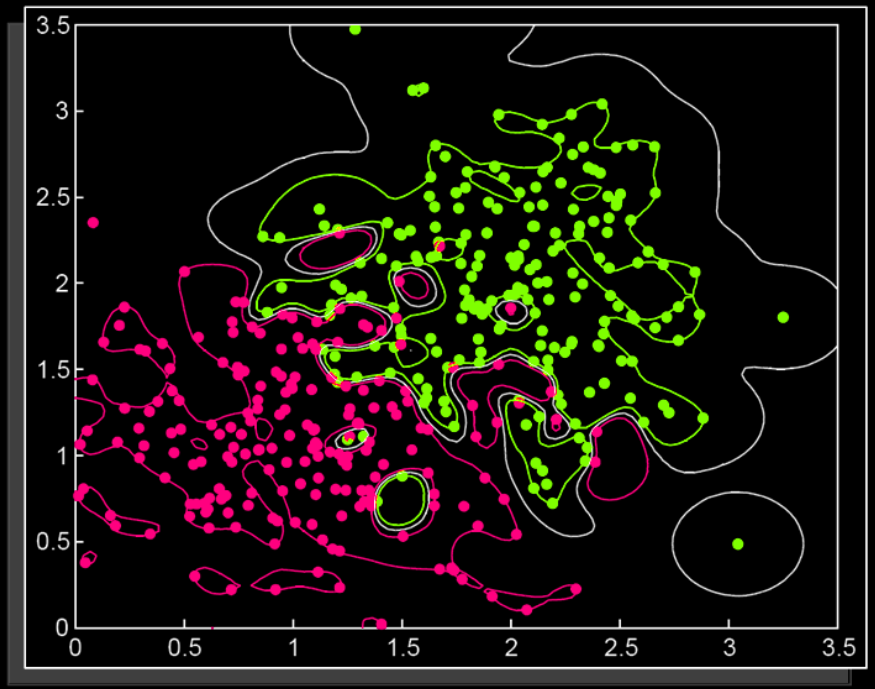
Descriptor space : no linear separation between data points.

Feature space : linear separation exists.

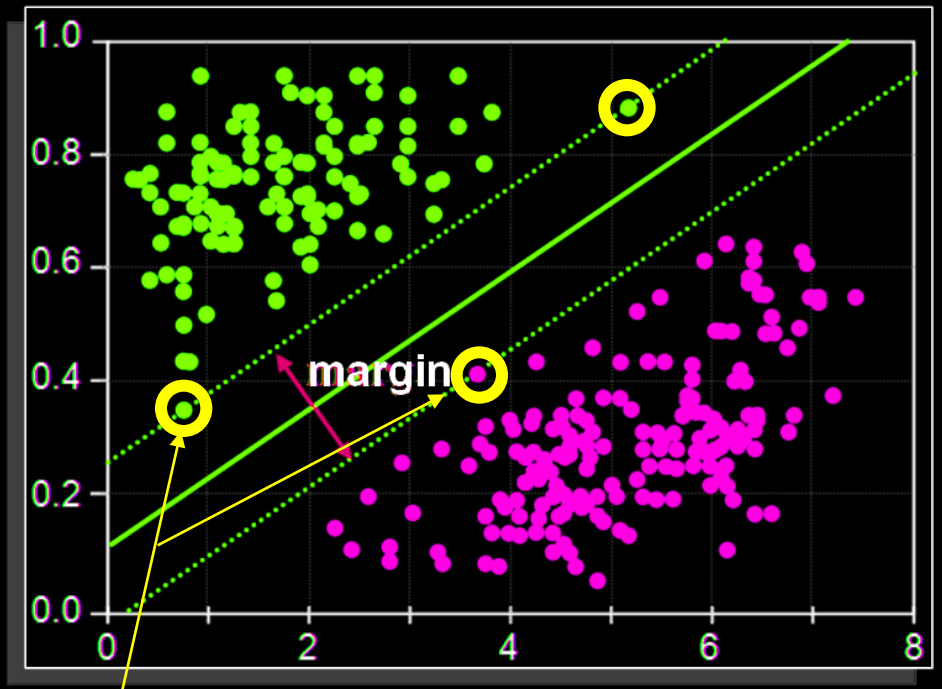
Support Vector Machine (SVM)

The SVM algorithm tend to maximize the margin around the hyperplane separating the two class of compounds. Different kernel functions and parameters have to be optimized (grid search) in order to identify the best models.

Descriptor space

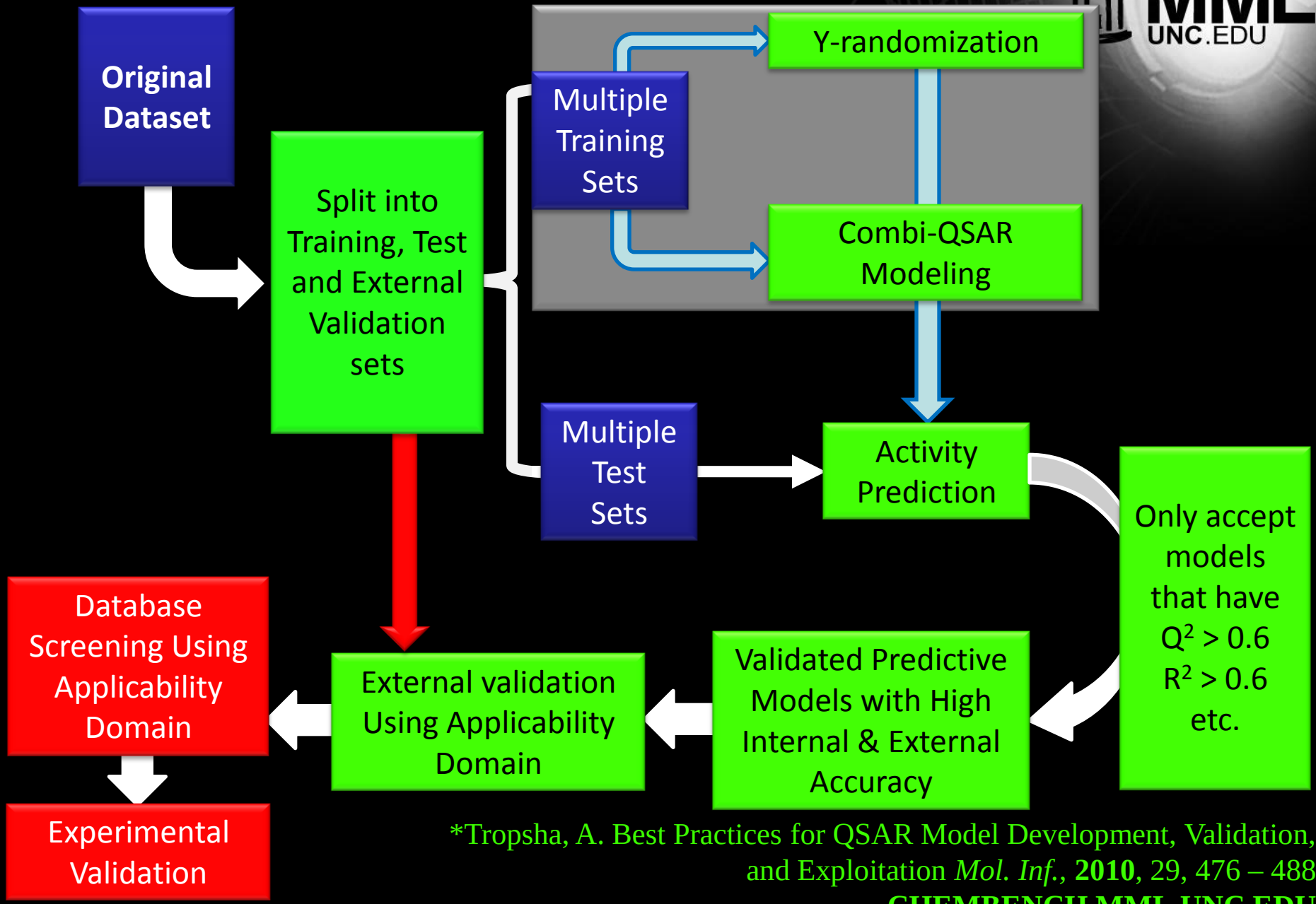


Feature space



Support vectors

Predictive QSAR Modeling Workflow*



*Tropsha, A. Best Practices for QSAR Model Development, Validation, and Exploitation *Mol. Inf.*, **2010**, 29, 476 – 488

CHEMBENCH.MML.UNC.EDU

Outline: From data to models to computer-aided design

MML
UNC.EDU

1.Data Accumulation: Need to establish an integrated database of nanomaterials' structure and properties

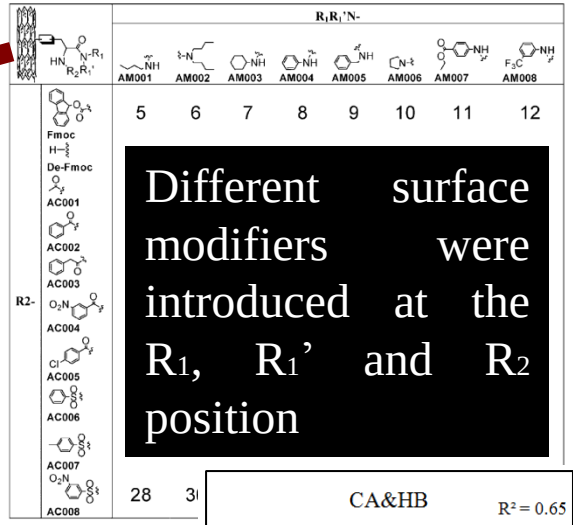
2. Data modeling and nano-informatics: Need to develop new descriptors for nanomaterials

3. Application of the Quantitative Structure Activity Relationships (QSAR) approach to modeling of nanomaterials

4. Proof-of-concept: Computer aided design of carbon nanotubes with the desired bioactivity and safety profiles

Case Study 3: Modeling of NPs for Protein Binding

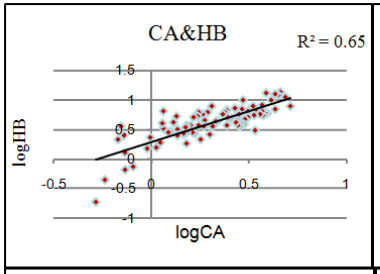
In 2008, Zhou et al* published *in vitro* protein binding, acute toxicity and immune toxicity assays for 84 Carbon NanoTubes (CNTs) decorated with different surface modifiers.

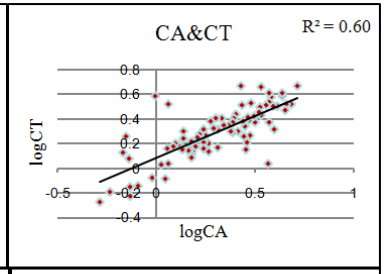


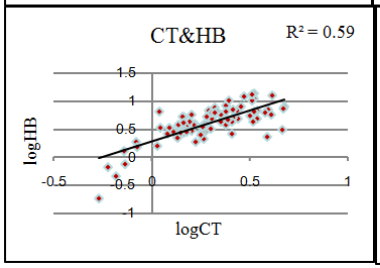
Different surface modifiers were introduced at the R_1, R_1' and R_2 position

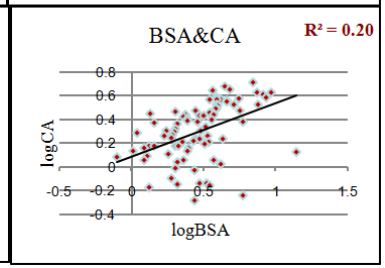
84 CNTs Tested in Two Different Types of Assays

- Protein Binding
  - Bovine Serum Albumin — BSA
  - Carbonic Anhydrase — CA
  - Chymotripsin — CT
  - Hemoglobin — HB
- Toxicity
  - Acute Toxicity
  - Immune Toxicity





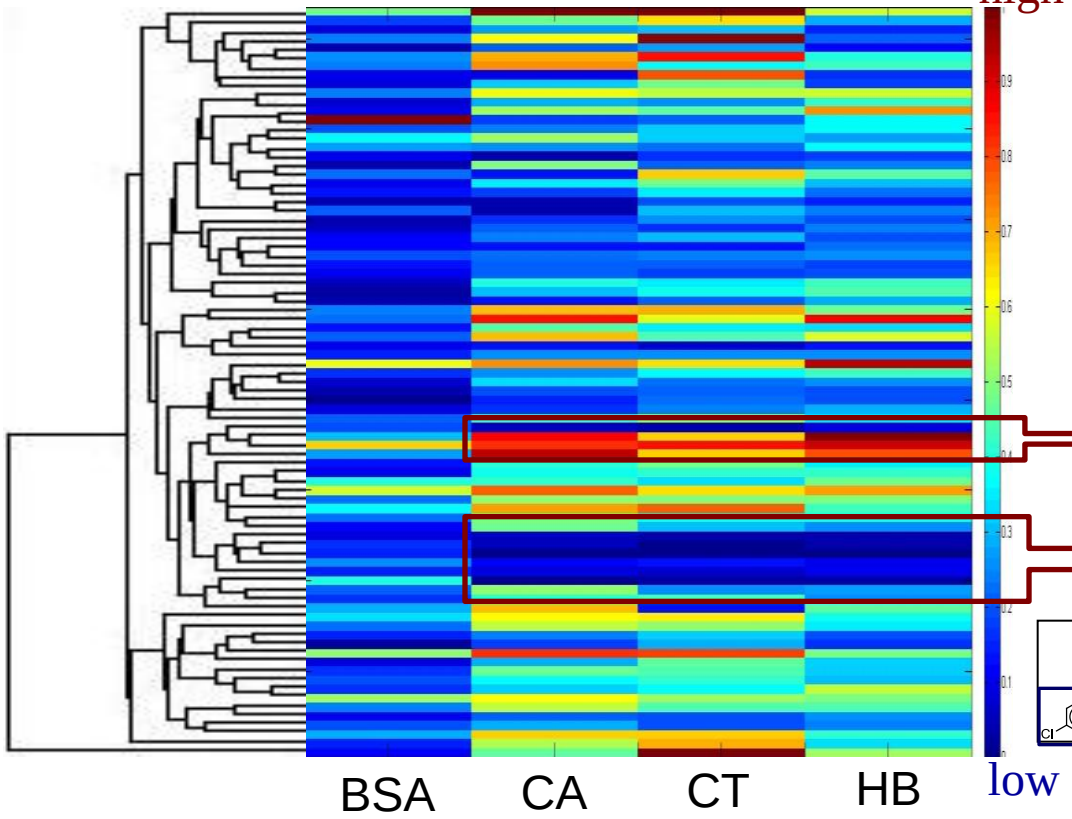




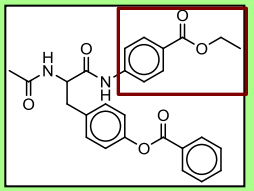
Zhou et al. Nano Lett., Vol. 8, No. 3, 2008

Case Study 3: Binding profiles sorted according to non-supervised hierarchical clustering of 84 NPs using chemical descriptors

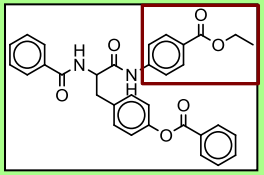
Clustering Uncovers Common Fragments with Distinct Protein Binding Profile.



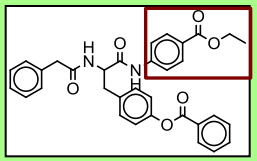
AMOO7AC001



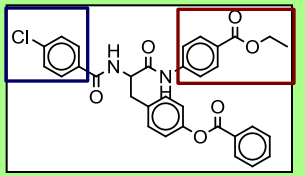
AMOO7AC002



AMOO7AC003



AMOO7AC005

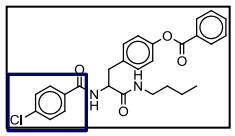


NP ID	BSA Binding	CA Binding	CT Binding	HB Binding
AMOO7AC001	4.0302	4.9276	3.5868	3.3936
AMOO7AC002	4.4012	4.6517	3.5818	4.2787
AMOO7AC003	8.5565	4.4131	4.4598	3.9182
AMOO7AC005	3.0478	0.7487	0.7748	0.3953

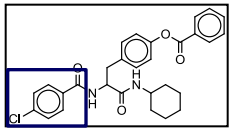
Cluster1: high binding affinities

Cluster2: low binding affinities

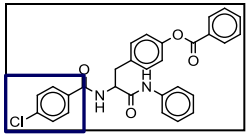
AMOO1AC005



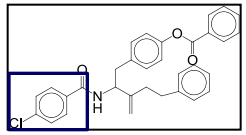
AMOO3AC005



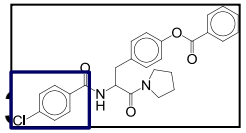
AMOO4AC005



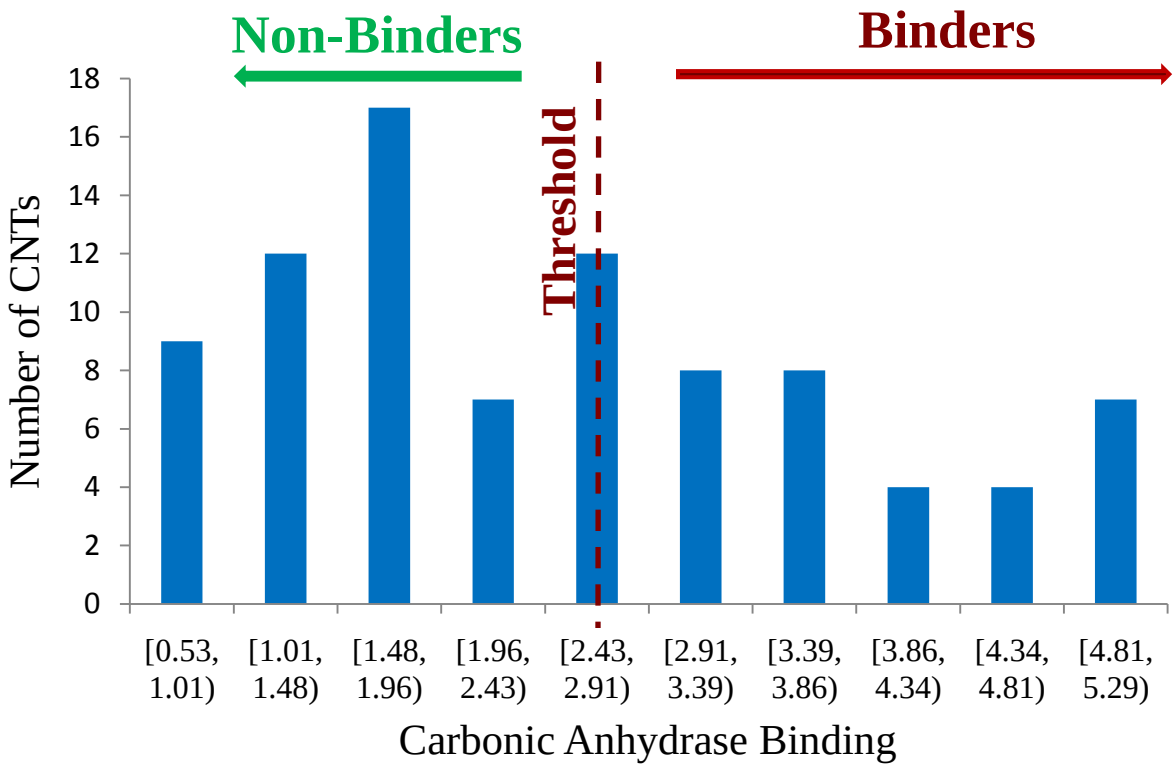
AMOO5AC005



AMOO6AC005



Case Study 3: QNAR Modeling of Carbonic Anhydrase Binding



Each CNT is represented by a single copy of its surface molecule.

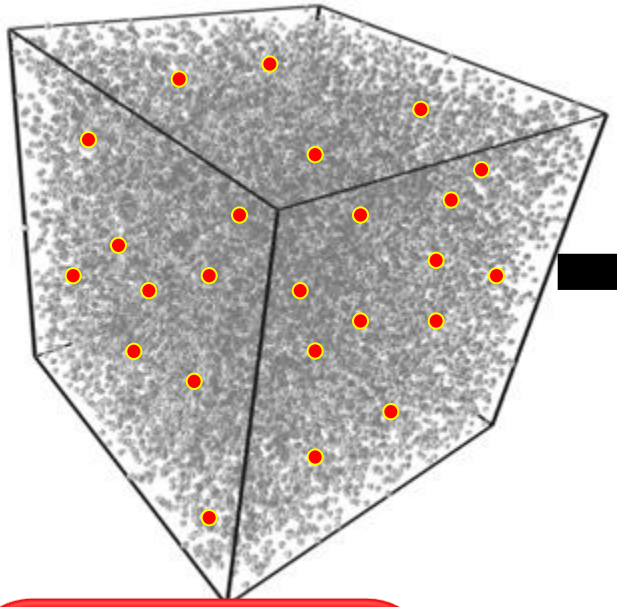
Consensus modeling approach combining different machine learning methods (k Nearest Neighbors, Support Vector Machines and Random Forest) and different types of chemical descriptors (Dragon and MOE).

Case study 3: QNAR Modeling of CA Binding

		kNN-Dragon	SVM-Dragon	RF-Dragon	kNN-MOE	SVM-MOE	RF-MOE
F1	Sens.	0.70	0.70	0.70	0.70	0.70	0.70
	Spec.	0.83	0.83	0.83	0.83	0.67	0.83
	Accr.	0.75	0.75	0.75	0.75	0.69	0.75
F2	Sens.	0.80	0.60	0.80	0.80	0.70	0.80
	Spec.	1.00	1.00	1.00	1.00	0.67	1.00
	Accr.	0.88	0.75	0.88	0.88	0.69	0.88
F3	Sens.	0.88	0.75	0.75	0.50	0.63	0.75
	Spec.	0.63	0.44	0.75	0.63	0.50	0.50
	Accr.	0.75	0.63	0.75	0.56	0.56	0.63
F4	Sens.	0.86	0.86	0.86	0.86	0.86	0.43
	Spec.	0.67	0.56	0.67	0.67	0.44	0.67
	Accr.	0.75	0.69	0.75	0.75	0.63	0.56
F5	Sens.	0.63	0.63	0.63	0.63	0.50	0.63
	Spec.	0.64	0.64	0.55	0.45	0.64	0.55
	Accr.	0.63	0.63	0.58	0.53	0.58	0.58
Total	Sens.	**0.77**	0.70	**0.74**	0.70	0.67	0.67
	Spec.	**0.73**	0.68	**0.73**	0.68	0.58	0.68
	Accr.	**0.75**	0.69	**0.73**	0.69	0.63	0.67

Case Study 3: Computer-aided design of novel carbon nanotubes with desired biological properties

(in collaboration with Dr. Bing Yan, St. Jude Children's Research Hospital)



Similarity Search

Consensus QSAR models

VIRTUAL SCREENING

~10^2 molecules

240,000 *in silico* designed small molecules which are considered attachable to CNTs

Similarity Search

462 molecules

QNAR Predictions

CA non-binders

110 molecules

Non-toxic

189 molecules

Experimental Validation

Experimental Validation (Toxicity assay)

MML UNC.EDU

Hits that were predicted as non-toxic

ID	Cell Viability (%)= 100 × treatment / control (percentage)				Average	STDEV
	Rep1	*Rep2*	*Rep3*	*Rep4*		
II-1 (1831)	55	53	60	63	58	5
II-6 (153907)	79	65	61	56	65	10
II-7 (153852)	73	72	66	62	68	6
II-8 (39260)	67	67	70	82	72	7
II-9 (13860)	72	67	66	67	68	3
II-10 (48636)	54	53	63	65	59	6

All rationally prioritized, synthesized, and tested CNTs predicted as non-toxic were confirmed experimentally.

Experimental Validation (Toxicity assay)

MML UNC.EDU

Hits that were predicted as toxic

ID	Cell Viability (%)= 100 × treatment / control (percentage)				Average	STDEV
	Rep1	*Rep2*	*Rep3*	*Rep4*		
II-11 (170243)	38	29	38	52	39	9
II-12 (170217)	49	50	38	58	49	8
II-16 (154018)	44	49	47	24	41	11
II-17 (4218)	44	51	54	56	51	5
II-18 (135817)	41	44	53	60	49	9
II-19 (120618)	47	43	66	62	55	11
II-20 (135018)	40	43	59	60	50	10

6 out of 10 rationally prioritized, synthesized, and tested CNTs predicted as toxic were confirmed experimentally

Conclusions

- Our results demonstrate that QNAR models can successfully predict the biological effects of MNPs from their descriptors either experimentally measured or calculated.

- Selected nanotubes decorated by the ligands identified with the help of QNAR models were experimentally synthesized and successfully validated. This study reports the first case of a rational design of carbon nanotubes possessing desired properties.

- QNAR models can be used to design new MNPs with appropriate bioactivity and safety profiles.

Needs

- *(side note: the technology to build models is in place: 50 years of QSAR!)*
- Sufficiently large datasets

  Both the core and the surface are varied systematically to dissect their relative effects on biology.
- NP-specific descriptors
  - easy for surface modifiers (unless there are non-linear core/surface modifier effects);
  - non-existent for the core (except maybe, QM calculations and MD-simulations derived?)
  - QSPR should precede QNAR
- Joint (e.g., virtual collaboratory) projects between computational and experimental teams

More questions and challenges than answers and solutions *(good for science*!)

- Relationships between *in vitro* and *in vivo* responses?
- Interplay between the structure of the core and that of surface modifiers as it affects toxicity?
- Greater understanding of the relationship between structure and physical properties of NPs (to impute the latter)
- Much greater effort is needed to generate designed datasets for focused QNAR investigations.
- Urgent need to develop ontology and integrated databases of structure and physical and biological properties of NPs

The Laboratory for Molecular Modeling

Principal Investigator
Alexander Tropsha

Research Professors
Clark Jeffries, Alexander Golbraikh, **Denis Fourches**

Graduate Research Assistants
Liying Zhang, Guiyu Zhao, Andrew Fant, Stephen Bush, Yen Low

Postdoctoral Fellows
Eugene Muratov, Aleks Sedykh, Ashutosh Tripathy, Nancy Baker

MML
UNC.EDU

MAJOR FUNDING

NIH
- R01-GM66940*
- R01-GM068665

EPA (STAR awards)
- RD832720
- RD833825